

Elektroakustische Nachbildung individueller Vokalklänge als Mittel für Sprachuntersuchungen

von KARL WILLY WAGNER, Friedrichsdorf i. Taunus

In einer früheren Arbeit¹⁾ habe ich ein elektrisches Sprechgerät zur Nachbildung der menschlichen Vokale beschrieben und eine Reihe von Versuchsergebnissen mitgeteilt. In den Schlussbemerkungen dieser Arbeit wurde darauf hingewiesen, dass das angegebene Verfahren nicht nur für den Physiker von Interesse ist, sondern auch dem Phonetiker, dem Sprachforscher, dem Physiologen und dem Mediziner wesentliche Dienste leisten kann. Einer freundlichen Anregung des Redaktors der «Folia phoniatica», den Gegenstand dem Leserkreis dieser Zeitschrift in einem kurzen Aufsatz näherzubringen, komme ich um so lieber nach, als ich seit 1936 an der Weiterarbeit gehindert wurde und meine Apparate im Krieg zerstört worden sind. Auch jetzt sehe ich noch keine Möglichkeit, an den Untersuchungen über die Sprachlaute weiterzuarbeiten; daher würde es mich freuen, wenn die Arbeit von anderer Seite fortgesetzt würde.

1. Entstehung und Beschaffenheit der natürlichen Vokalklänge

Die vokalartigen Klänge entstehen in der Weise, dass die Luft

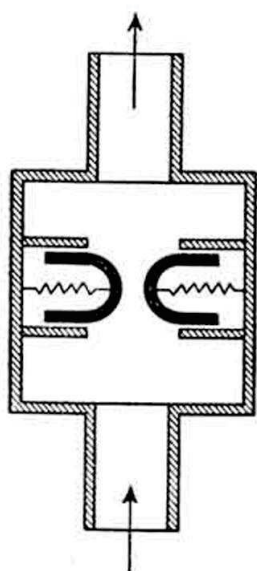


Abb. 1. Prinzip der Eucaldschen Polsterpfeife.

1) K. W. Wagner, «Ein neues elektrisches Sprechgerät zur Nachbildung der menschlichen Vokale», Abh. Nr. 3 der Preuss. Akademie der Wissenschaften 1936.

in den verschiedenen Hohlräumen des Sprechapparates — Luftröhre, Rachenraum, Mundhöhle, Nasenhöhle — durch den Kehlkopf zu Schwingungen angeregt wird. Beim Singen oder Sprechen werden die Stimmbänder aneinander gepresst und durch die Kehlkopfmuskulatur mehr oder weniger gespannt, so dass die ausgeatmete Luft nur stossweise austreten kann. Durch diese Luftstösse werden die verschiedenen Tonräume zu Eigenschwingungen angeregt. Die Frequenz der im Kehlkopf erzeugten Luftstösse wird durch die Muskelspannung eingestellt; sie bestimmt den Grundton des Klanges. Man hat früher die Funktion des Kehlkopfes mit der einer Zungenpfeife verglichen; sie entspricht aber mehr der von *J. R. Ewald*²⁾ angegebenen Polsterpfeife, die in Abb. 1 schematisch dargestellt ist. Eine der Wirkungsweise des Kehlkopfs gut angepasste Ausführungsform der Polsterpfeife hat *F. Wethlo* konstruiert und eingehend untersucht³⁾. In mehreren schönen Arbeiten von *W. Trendelenburg*⁴⁾ ist der Mechanismus des Kehlkopfes und die Art der erzeugten Luftstösse in vielen Einzelheiten weiter aufgeklärt worden.

Wie schon erwähnt, sind bei der Erzeugung der Sprachklänge mehrere Tonräume beteiligt, die durch ihre gegenseitige Verbindung miteinander gekoppelt sind. Gestalt und Kopplung der Tonräume kann durch die Stellung der Sprachorgane weitgehend verändert werden; dabei fällt besonders der Zunge eine wichtige Rolle zu. Durch diese Variationen kommt die Fülle der verschiedenen Sprachlaute zustande.

Physikalisch kann man sich den Vorgang der Lauterzeugung folgendermassen vorstellen. Im Kehlkopf wird eine regelmässige Folge von Luftstössen hervorgebracht; den Intensitätsverlauf möge die Abb. 2 veranschaulichen. Wenn die Pause zwischen zwei aufeinanderfolgenden Stössen im Vergleich zur Stossdauer hinreichend lang ist, kommt es auf die Form des Stosses nicht an;

$$f = \frac{1}{T}$$

ist die Grundfrequenz des Tones.

2) *J. R. Ewald*, Pflügers Arch. f. d. ges. Physiol., Bd. 152 (1913), S. 171.

3) *F. Wethlo*, «Versuche mit Polsterpfeifen», Beitr. z. Anatomie, Physiol. Pathol. und Therapie des Ohres, der Nase und des Halses, Bd. 6 (1913), S. 268.

4) *W. Trendelenburg*, in einer Reihe von Abhandlungen der Preussischen Akademie der Wissenschaften seit 1936.

Die Gesamtheit der Tonräume im menschlichen Sprechapparat stellt ein gekoppeltes Schwingungssystem dar. Es besitzt eine Reihe von Eigenschwingungen; ihre Eigenfrequenzen seien

$$f_1, f_2, f_3 \dots$$

Die Dämpfungen dieser Eigenschwingungen bezeichnen wir mit

$$d_1, d_2, d_3 \dots$$

Jeder Luftstoss regt diese Schwingungen an, so dass eine zusammengesetzte Schwingung der Form

$$\varphi(t) = A_1 e^{-d_1 t} \sin(2\pi f_1 t + \varphi_1) + A_2 e^{-d_2 t} \sin(2\pi f_2 t + \varphi_2) + \dots$$

entsteht; die $A_1, A_2 \dots$ sind die Amplituden der einzelnen Eigenschwingungen. Es handelt sich, wie man sieht, um eine ziemlich komplizierte Schwingungsform. Da die Luftstösse periodisch aufeinander folgen, wiederholt sich auch die Schwingung nach Ablauf der Periodendauer T immer von neuem. Im Endergebnis entsteht eine rein periodische Schwingung mit der Grundfrequenz f und zahlreichen *harmonischen Obertönen* mit den Frequenzen

$$2f, 3f, 4f \dots$$

Dass die Obertöne harmonisch werden, obgleich die Eigenfrequenzen $f_1, f_2 \dots$ des gekoppelten Systems in der Regel unharmonisch sind, kann man sich durch die folgende Ueberlegung klar



Abb. 2. Zeitlicher Verlauf der im Kehlkopf hervorgebrachten Luftstösse.

machen. Die periodische Stossfolge nach Abb. 2 kann man nach *Fourier* in ein Spektrum reiner Sinusschwingungen auflösen. Es enthält die Frequenzen $f, 2f, 3f \dots$. Die Amplituden sind praktisch konstant bis zur Ordnungszahl n , wobei ungefähr

$$\frac{1}{nf} = \tau$$

ist, wo τ die Stossdauer bedeutet; Abb. 3 stellt im oberen Bild ein solches Spektrum dar.

Wirkt ein derartiger Komplex von Schwingungen auf eine Anordnung mit den Eigenfrequenzen $f_1, f_2, \dots$ so werden diejenigen Schwingungskomponenten besonders stark hervorgehoben, die den

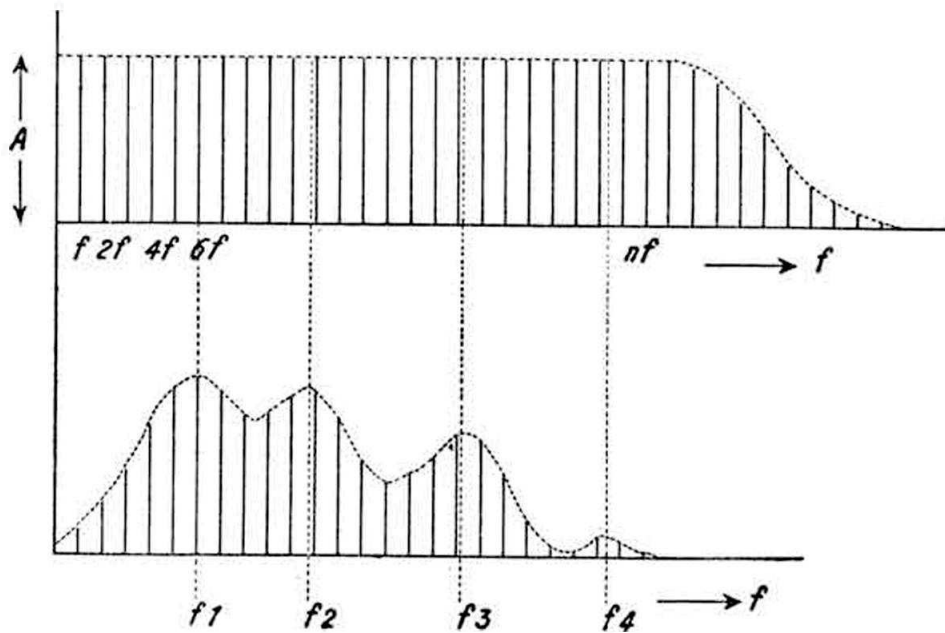


Abb. 3. oben: Tonspektrum der vom Kehlkopf erzeugten periodischen Folge von Luftstößen.
 unten: Tonspektrum der Schwingungen, die im Sprechorgan durch den Kehlkopf angeregt werden.

Eigenfrequenzen benachbart sind. Es entsteht ein Schwingungszustand mit einem Spektrum nach Art des in Abb. 3 unten dargestellten. Ändert man die Grundfrequenz f , so verschieben sich zwar die Linien des Spektrums entsprechend, die gestrichelt gezeichnete Umhüllende des Spektrums bleibt jedoch erhalten, weil die Lage der Eigenfrequenzen $f_1, f_2, \dots$ und die Grösse ihrer Dämpfungen unverändert bleibt; denn diese Bestimmungsstücke hängen nur von der Gestaltung der Tonräume und von ihrer Kopplung ab, nicht aber von der Art der Anregung.

Spektren der in Abb. 3 unten angegebenen Art werden in der Tat erhalten, wenn man die Schwingungskurven von Vokalklängen in ihre harmonischen Teilschwingungen zerlegt; entweder durch die *Fourier'sche* Analyse, oder mit irgend einem der bekannten elektrischen Interferenzverfahren mit gleitendem Suchton, etwa nach *M. Grützmacher*⁵⁾. Solche Zerlegungen von Vokalklängen in ihre harmonischen Komponenten sind planmässig und umfassend be-

⁵⁾ *M. Grützmacher*, „Eine neue Methode der Klanganalyse“, *Elektr. Nachr. Techn. (ENT)*, Bd. 4 (1927), S. 533.

sonders von *C. Stumpf*⁶⁾, *D. C. Miller*⁷⁾ und *A. Gemelli* und *G. Pastori*⁸⁾ durchgeführt worden. Sie alle bestätigen die v. *Helmholtz*'sche Vokaltheorie⁹⁾, wonach

1. dem stationären Teil des Vokalklages¹⁰⁾ eine periodische Schwingung entspricht;

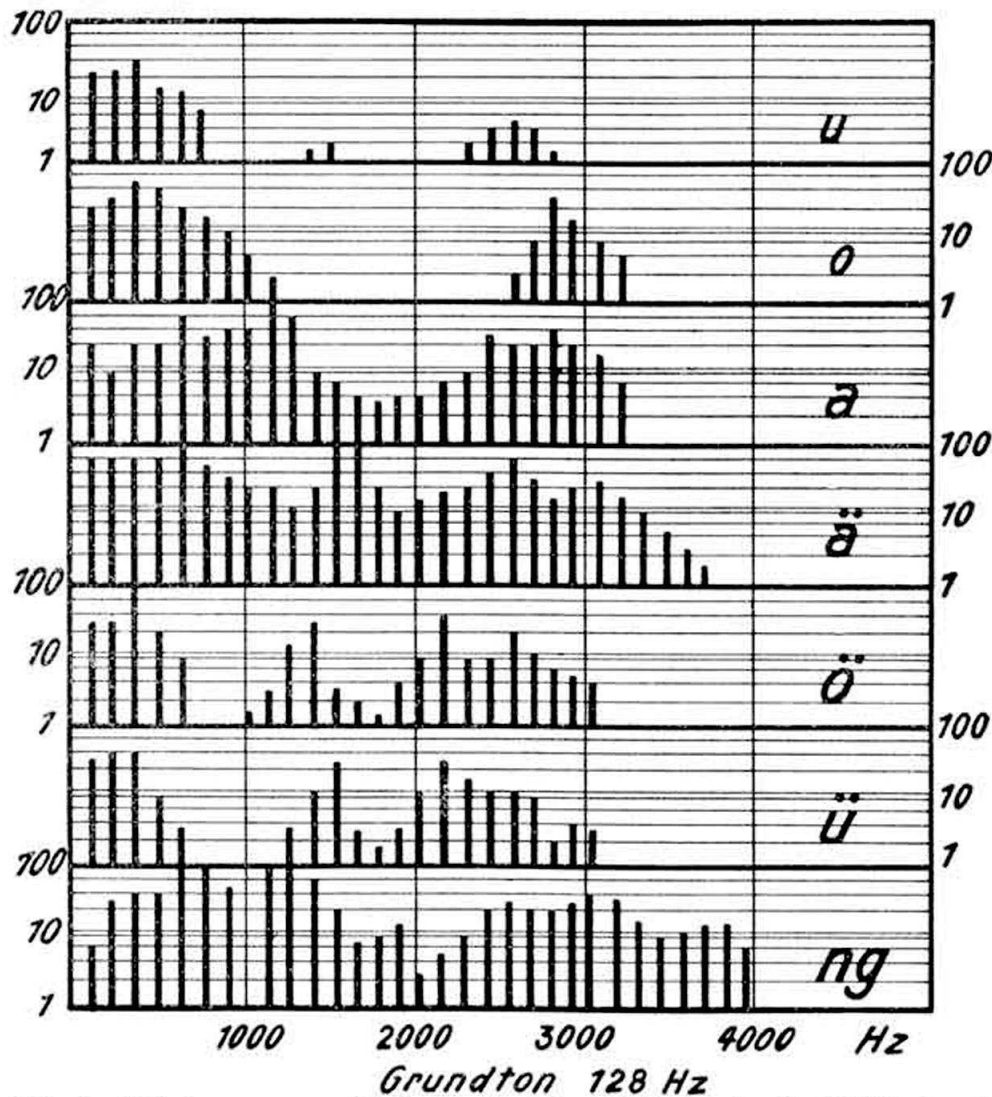


Abb. 4. Teiltöne von verschiedenen deutschen Vokalen (nach E. Thienhaus).

⁶⁾ *C. Stumpf*, Sitz. Ber. d. Preuss. Akad. d. Wiss. vom 4.4. 1918. «Die Sprachlaute», Berlin 1926.

⁷⁾ *D. C. Miller*, «The Science of Musical Sounds», New York 1922.

⁸⁾ *A. Gemelli* und *G. Pastori*, «L'Analisi elettroacustica del Linguaggio», Mailand 1934.

⁹⁾ *H. v. Helmholtz*, «Die Lehre von den Tonempfindungen», 6. Ausg., S. 168 ff., Braunschweig 1913.

¹⁰⁾ Der An- und Ausklang wurde besonders von *Gemelli* und *Pastori* untersucht; siehe Fussnote 8.

2. der besondere Vokalcharakter durch das Hervortreten der Teiltöne in bestimmten Frequenzbereichen gegeben ist. Ihre Lage ist für jeden Vokal charakteristisch und praktisch unabhängig von der Höhe des Grundtons, auf dem der Vokal gesungen oder gesprochen wird. Man nennt diese Frequenzbereiche die «*Formantgebiete*» oder kurz die *Formanten* des betreffenden Vokals.

Nach der vorher dargelegten Entstehungsweise der Sprachlaute sind die Formantfrequenzen die Eigenfrequenzen des Sprechapparats.

Abb. 4 zeigt die von E. Thienhaus¹¹⁾ mittels Suchtonverfahrens bestimmten Spektren von verschiedenen Vokalen der deutschen Sprache. Man erkennt, dass neben den beiden schon lange bekannten Hauptformanten bei fast allen Vokalen noch weitere, weniger hervortretende Verstärkungsgebiete von Obertönen auftreten; C. Stumpf hat sie *Nebenformanten* genannt. Die genaue Lage der verschiedenen Formanten und ihr Stärkeverhältnis bestimmt den individuellen Charakter des Vokalklangs; wir kommen hierauf später zurück.

2. Künstliche Vokalklänge

Die Bemühungen um die Erforschung der Vokalklänge und um ihre künstliche Wiedergabe sind alt. Schon im Jahre 1779 erliess die Akademie der Wissenschaften in Petersburg ein Preisausschreiben mit den beiden folgenden Fragen:

- a) Worin besteht das Wesen und das Kennzeichen der verschiedenen Vokalklänge?
- b) Ist es möglich, einen Apparat zur Erzeugung der Vokalklänge zu bauen, der diese genau wiedergibt?

Den Preis erhielt Prof. Kratzenstein für eine Reihe von resonatorartigen Gebilden, die durch den von einer Zunge unterbrochenen Luftstrom erregt wurden. Sir David Brewster¹²⁾ berichtet von diesen Apparaten, dass sie die fünf Hauptvokale gut erkennbar wiedergaben. Ähnliche Versuche beschrieb R. Willis¹³⁾ 1829; er benutzte Zungenpfeifen mit zylindrischen Ansatzrohren, die auf den Haupt-

¹¹⁾ E. Thienhaus, «Neuere Versuche zur Klangfarbe und Lautstärke von Vokalen», Z. f. techn. Phys., Bd. 15 (1934), S. 637. L. Barczinski und E. Thienhaus, «Klangspektren und Lautstärke deutscher Sprachlaute», Arch. Néerl. de Phonétique Expérimentale; Bd. 11 (1935), S. 47.

¹²⁾ Sir David Brewster, «Letters on natural magic...», 1832, S. 207.

¹³⁾ R. Willis, Trans. of Cambr. Phil. Soc., Bd. 3 (1829), S. 231; Pogg. Ann. d. Phys., Bd. 24 (1832), S. 307.

formanten abgestimmt waren, oder erregte die Schwingung in diesen Röhren mittels Blattfeder und Zahnrad. Besonders gut gelang die Erzeugung der dunklen Vokale; die auf diese Weise ermittelten Formantfrequenzen sind erstaunlich genau. *H. v. Helmholtz* hat die Vokale *a*, *o*, *u* mit Hilfe von elektrisch angeregten Stimmgabeln erzeugt. *D. C. Miller*⁷⁾ bildete die Vokale durch Sätze von Orgelpfeifen nach; für das kurze *ä* (englisches Wort *mat*) brauchte er 16 Pfeifen und gibt an, dass selbst hiermit die Nachbildung zu wünschen übrig liess. Vokalsynthesen von grosser Naturtreue gelangen *C. Stumpf*⁶⁾ mittels 28 Lippenpfeifen, die durch Interferenzröhren von allen Obertönen gereinigt waren. *F. Wethlo*³⁾ erzielte gute Vokalklänge durch Anblasen von Resonatoren mittels Polsterpfeife. Noch wesentlich weiter ist *R. Paget*¹⁴⁾ in der Nachbildung des menschlichen Sprechapparates gegangen. Er formte aus Plastilin Modelle der Mundhöhle und des Rachenraumes und erregte sie durch einen künstlichen Kehlkopf. Auf diese Weise konnte er alle englischen Vokale künstlich herstellen.

Auch auf elektrischem Wege kann man Vokalschwingungen erzeugen und durch Lautsprecher hörbar machen. In den von *J. O. Stewart*¹⁵⁾ *H. Fletcher*¹⁶⁾ und *F. Trautwein*¹⁷⁾ angegebenen Schaltungen dient als elektrischer Kehlkopf ein obertonreicher Schwingungsgenerator; er erregt einen oder mehrere elektrische Schwingungskreise in Reihen- oder Parallelschaltung.

Wenn man nach *R. Paget* den natürlichen Sprechapparat in allen Einzelheiten nachbildet, so muss die künstlich erzeugte Vokalschwingung der natürlichen genau entsprechen. Abgesehen von der in solchen Versuchen offenbarten technischen Meisterschaft können sie aber nichts Neues bieten. Wesentlich aufschlussreicher sind die Versuche von *Miller* und *Stumpf*, sowie die elektrisch erzeugten Vokalklänge; denn sie erbringen den Nachweis, dass die Vokaltheorie von *Helmholtz* die wesentlichen Kennzeichen des Vokalklanges umfasst und beweisen die Richtigkeit des *Helmholtz'schen* Satzes, der die Unabhängigkeit des Klangeindrucks von der Phasenlage der Teil-

¹⁴⁾ *R. Paget*, Proc. Roy. Soc. (A) London, Bd. 102 (1923), S. 752; Bd. 106 (1924), S. 150; «Human Speech», London 1930.

¹⁵⁾ *J. O. Stewart*, «An electrical analogue of the vocal organs», Nature, Bd. 110 (1922), S. 311.

¹⁶⁾ *H. Fletcher*, «Speech and Hearing», New York und London 1929.

¹⁷⁾ *F. Trautwein*, «Elektrische Musik», Berlin 1930.

töne zueinander behauptet. Denn bei den genannten Anordnungen ist die Phasenlage eine rein zufällige.

3. Ein elektrisches Gerät für die Nachbildung individueller Vokale

Zur Erforschung der Kennzeichen für die individuelle Verschiedenheit von Vokalklängen empfiehlt sich besonders das Verfahren der elektrischen Nachbildung. Man kann es so einrichten, dass die Amplituden, die Frequenzen und die Dämpfungen der einzelnen Formanten bequem eingestellt und schnell geändert werden können, so dass man die Einflüsse all dieser Grössen auf den Vokalklang leicht studieren kann. Der von mir gebaute Apparat erfüllt diese Forderung in vollem Masse. Da er in der unter ¹. genannten Veröffentlichung ausführlich beschrieben ist, kann ich mich hier auf eine kurze Darlegung der allgemeinen Anordnung beschränken. Sie ist in Abb. 5 schematisch wiedergegeben.

Die links mit S.O. und E bezeichnete Einrichtung ist ein Oszillator zur Erzeugung von periodischen Stromstössen nach Abb. 2.

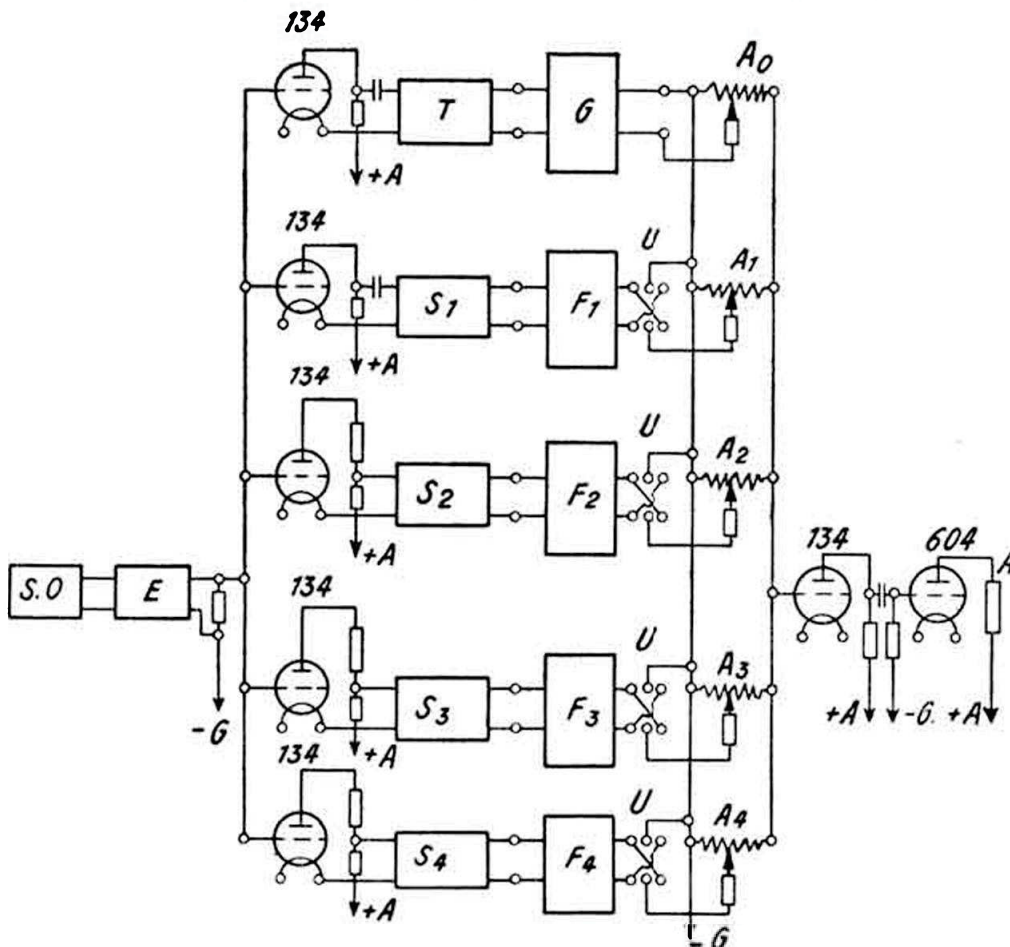


Abb. 5. Schaltung des elektrischen Geräts zur Nachbildung von Vokalen.

Es ist der «elektrische Kehlkopf»; wir nennen ihn kurz Stosskreis oder Stossoszillator.

Die Nachbildung des Sprechorgans enthält vier Formantkreise F_1 bis F_4 mit einstellbarer Eigenfrequenz und einstellbarer Dämpfung. Sie werden vom Stosskreis her über je eine Verstärkerröhre RE 134 und eine elektrische Siebkette (S_1 bis S_4) erregt. Es hat sich nämlich als zweckmässig erwiesen, nicht das ganze Frequenzspektrum des Stosskreises (Abb. 3, oben) auf den Formantkreis wirken zu lassen, sondern nur ein die Eigenfrequenz umgebendes, nicht gar zu breites Frequenzband; man erhält so leichter saubere Formantschwingungen. Diese werden über Spannungsteiler (A_1 bis A_4) und einen zweistufigen Verstärker (ein Rohr RE 134 und ein Rohr RE 604) dem Wiedergabeparaat A (Lautsprecher) zugeführt. Mittels der Spannungsteiler wird die Amplitude jeder Formantschwingung eingestellt.

Da die Siebketten S_1 bis S_4 den Grundton mehr oder weniger unterdrücken, muss er durch einen besonderen, auf den Grundton abgestimmten Kreis G dem Tongemisch wieder hinzugesetzt werden. Dem Kreis G ist ein Tiefpass T und auch ein Verstärkerrohr RE 134 vorgeschaltet. Mittels des Spannungsteilers A_0 wird die Grundtonamplitude eingestellt. In manchen Fällen übernimmt dieser Kreis auch die Bildung tiefliegender Formanten.

Durch die im Bild links gezeichneten Verstärkerröhren wird eine Rückwirkung der Formantschwingungen auf den Stosskreis verhindert. Ferner enthält jeder Formantkreis auf seiner Ausgangsseite gleichfalls eine (im Bild nicht gezeichnete) Verstärkerstufe; dadurch wird die gegenseitige Beeinflussung der Formantkreise unterbunden. Auf diese Weise lässt sich jeder Formantkreis beliebig verstellen, ohne dass dabei die Schwingungen der übrigen Kreise im geringsten beeinträchtigt werden. Ohne diese Massnahme wäre das Experimentieren mit der Anlage sehr erschwert.

An den für die akustische Wiedergabe der Schwingungen benutzten Lautsprecher werden hohe Anforderungen gestellt; er muss den ganzen in Betracht kommenden Frequenzbereich gleichmässig übertragen; er darf also nicht selbst Verzerrungen hervorrufen. Ferner muss der Wiedergaberaum gut gedämpft sein, weil Raumresonanzen ausserordentlich störend wirken¹⁸⁾.

¹⁸⁾ Beispiele für solche Verzerrungen sind in der unter 1. genannten Arbeit angegeben.

4. *Ergebnisse und Probleme*

Aus den sehr zahlreichen Versuchsreihen, die mit dem beschriebenen Gerät ausgeführt worden sind, kann ich hier nur einen ganz kleinen Ausschnitt bringen; dabei soll auch über die bei den Versuchen gewonnenen allgemeinen Erfahrungen berichtet werden.

Will man einen Vokalklang nachbilden, so ist es zweckmässig, mit einer Analyse des natürlichen Klanges zu beginnen; denn man braucht ja Anhaltspunkte für die Einstellung des Nachbildungsgerätes. Man kann diese natürlich auch nur nach dem Gehör ausführen; das ist aber ein sehr mühsames und langwieriges Verfahren, da die Zahl der Variablen (Amplitude, Frequenz und Dämpfung der Schwingungen in fünf verschiedenen Kreisen) gross ist. Ursprünglich wurde die Schwingungskurve des natürlichen Klanges oszillographisch aufgezeichnet und das Oszillogramm nach dem von *C. Runge* angegebenen Rechenschema unter Benutzung der von *P. Terebesi* herausgegebenen Schablonenblätter¹⁹⁾ analysiert. Das Verfahren ist zeitraubend und liefert auch nur die ersten 11 Teiltöne, was in vielen Fällen nicht ausreicht. Eine weitere Schwierigkeit besteht darin, dass bei den meisten Menschen der Kehlkopf etwas unregelmässig arbeitet und daher keine genau periodischen Luftströme liefert. Infolgedessen sind die aufeinander folgenden Perioden der Schwingungskurve des Vokalklanges zwar ähnlich, aber nicht gleich. Man muss deshalb eine Reihe von Perioden analysieren und aus den Ergebnissen den Mittelwert berechnen. Bequemer ist das elektrische Analysierverfahren mittels Suchtons, wobei der Mittelwert selbsttätig gebildet wird.

Wenn die Versuchsperson eine geschulte Stimme hat und musikalisch ist, geht die Analyse ohne weitere Umstände von statten. Die meisten Sprecher können aber während des Versuchs die Tonhöhe nicht genau einhalten, wodurch die Analysen unter Umständen erheblich gefälscht werden. Auch diese Schwierigkeit konnte schliesslich durch ein besonderes Versuchsverfahren¹⁾ vollständig überwunden werden.

Die aus der Analyse ermittelten Formantfrequenzen werden am Sprechgerät eingestellt. Man bekommt auf diese Weise bereits einen dem natürlichen Vokal ziemlich ähnlichen Klang. Nun beginnt die

¹⁹⁾ *P. Terebesi*, Rechenschablonen für harmonische Analyse nach *C. Runge*, Berlin 1930.

Feineinstellung der Amplituden und Dämpfungen der verschiedenen Formanten und des Grundtons nach dem Gehör; auch die Formantfrequenzen müssen manchmal nachgestellt werden. Auf diese Weise ist es gelungen, die individuellen Vokalklänge aller Versuchspersonen so gut nachzubilden, dass der natürliche und der künstlich erzeugte Klang nicht unterscheidbar waren. Dies bezieht sich auch auf die verschiedenen Varianten der Umlaute ä, ö, ü und die nasalen Vokale der französischen Sprache. Die Versuche wurden nicht nur im Laboratorium gemacht, sondern auch einem kleineren und einem grossen Zuhörerkreis vorgeführt²⁰).

Die nachstehenden Einzelergebnisse vermögen nur einen unvollkommenen Eindruck der Leistungsfähigkeit des Verfahrens zu vermitteln, weil der Hörvergleich fehlt und weil aus Raumgründen nur einige wenige Beispiele wiedergegeben werden können.

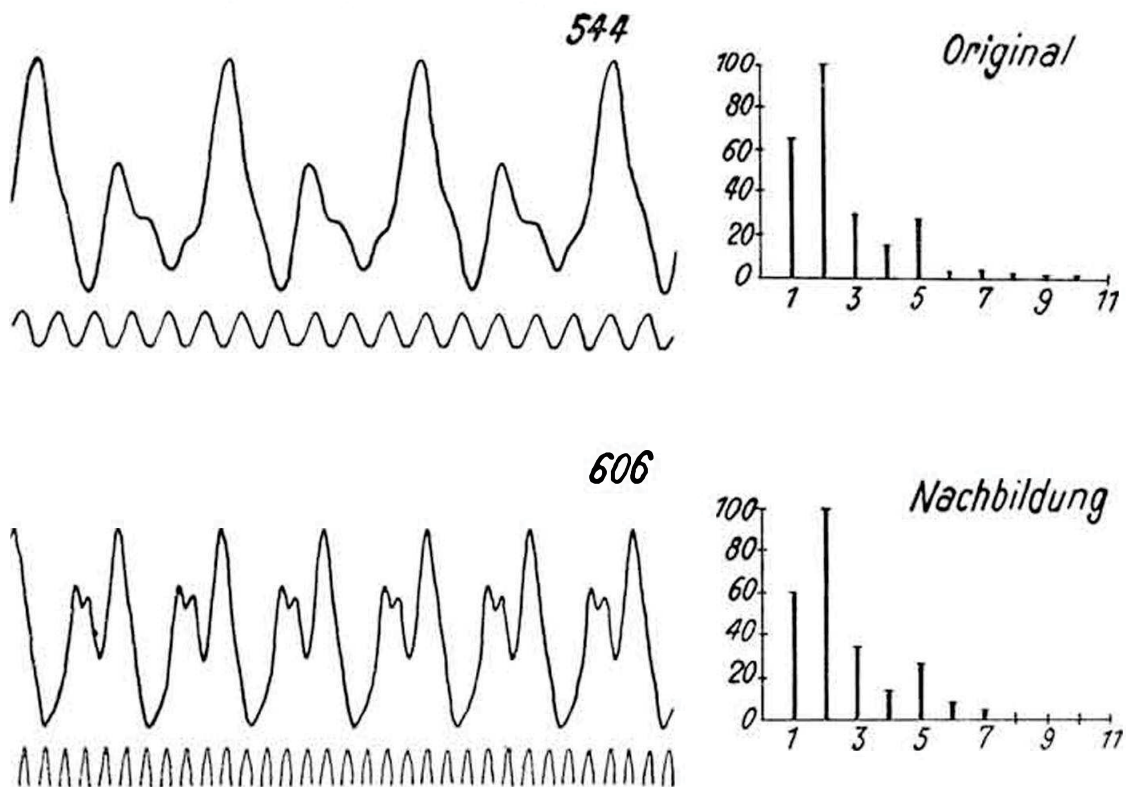


Abb. 6. Original und Nachbildung des Vokals u des Sprechers Wz,
Tonhöhe 156 Hz.

²⁰) Vorführung in der Gesamtsitzung der Berliner Akademie am 12. 1. 1933; in einer gemeinsamen Sitzung des Elektrotechnischen Vereins und der Heinrich-Hertz-Gesellschaft zur Förderung des Funkwesens im grossen Hörsaal d. Berliner Techn. Hochschule am 20. 2. 1934 wiederholt.

Abb. 6 zeigt links die Schwingungskurve und rechts das Teiltenspektrum des Vokals *u* eines Mitarbeiters (Wz) auf der Tonhöhe von 156 Hertz. Das obere Bild bezieht sich auf den natürlichen Klang, das untere Bild auf den künstlich erzeugten; beide sind mit demselben Kondensatormikrophon aufgenommen worden. Die Verschiedenheit der Schwingungskurven hat keine Bedeutung; die Form dieser Kurve hängt von den Nullphasenwinkeln der Teiltöne ab; die gegenseitige Phasenbeziehung hat jedoch keinen Einfluss auf den Gehöreindruck. Massgebend ist die weitgehende Uebereinstimmung der Teiltenspektren; ihr entsprach die Beobachtung, dass der künstlich erzeugte Vokal dem natürlichen vollständig gleich.

In den Abb. 7 und 8 sind die Nachbildungsergebnisse des Vokals *a* eines anderen Sprechers (H) auf den Tonhöhen 117 und 200 Hertz wiedergegeben. Bei dem Versuch zu Abb. 7 waren vier Schwingungskreise mit den folgenden Einstellungen eingeschaltet:

	f Hertz	d	A %
G	211	0,9	—
F_1	662	0,16	100
F_2	1140	0,21	82
F_3	1440	0,06	80

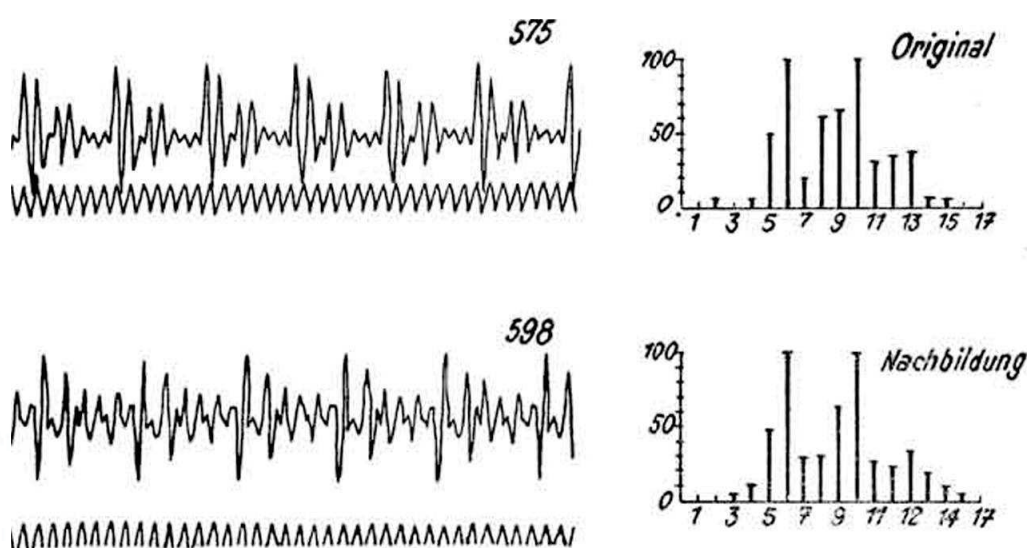


Abb. 7. Original und Nachbildung des Vokals *a* des Sprechers H, Tonhöhe 117 Hz.

Neben den beiden kräftigen Formanten mit den Tonhöhen 662 und 1140 Hertz tritt noch ein mittelstarker Oberformant bei 1440 Hertz und ein schwacher Unterformant bei 211 Hertz auf.

Wird der Grundton auf 200 Hertz erhöht (Abb. 8), so fällt der Formant G nahezu auf den Grundton; F_1 verschiebt sich nach 840 Hertz, F_2 bleibt fast unverändert bestehen; der Formant bei 1440 Hertz ist nicht mehr nachweisbar; dagegen tritt ein schwacher Nebenformant bei 1900 Hertz auf; siehe die Werte in der nachstehenden Tafel.

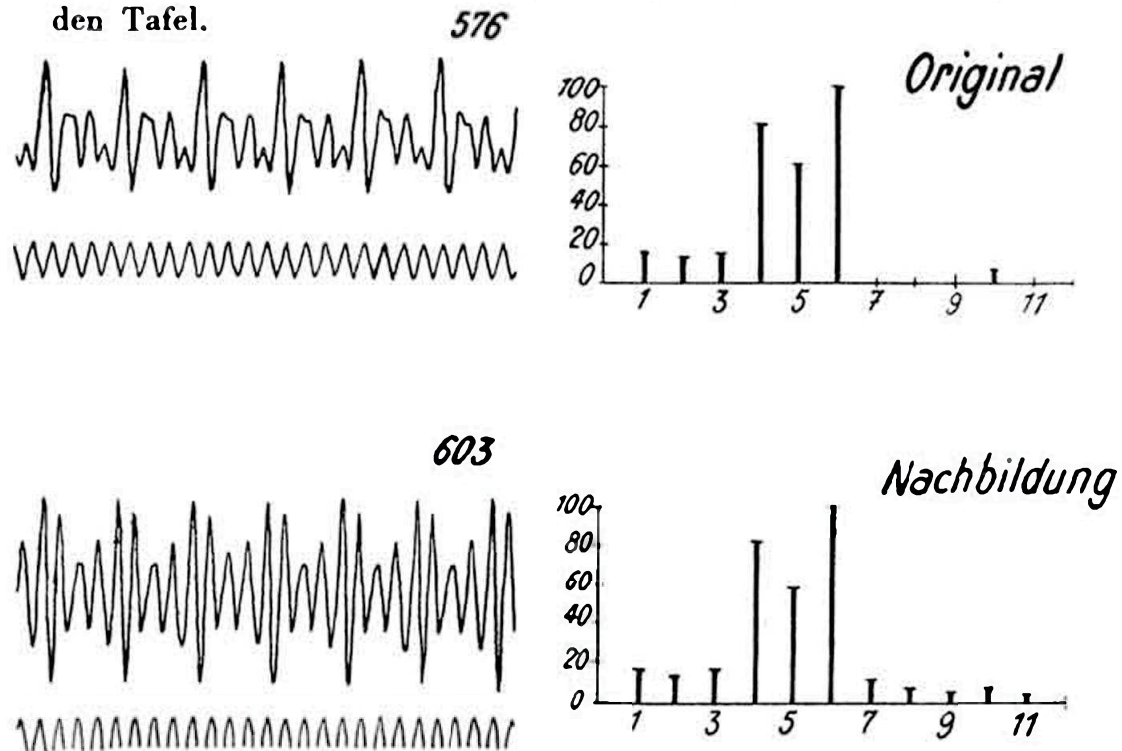


Abb. 8. Original und Nachbildung des Vokals a des Sprechers H, Tonhöhe 200 Hz.

	f Hertz	d	A %
G	263	1,6	1
F_1	840	0,3	100
F_2	1190	0,1	58
F_3	1900	0,08	6,8

Die teilweise Formantverschiebung ist darauf zurückzuführen, dass der Sprecher mit der Veränderung der Tonhöhe auch die Mundstellung etwas ändert; siehe die Darlegungen zu Abb. 11. Wird die

Mundstellung beibehalten und lediglich die Tonhöhe durch den Kehlkopf verändert, so bleiben die Formantfrequenzen bei allen Tonhöhen die gleichen. Dies konnte z. B. in der Weise gezeigt werden, dass der Sprecher (in diesem Falle ich selbst) denselben Vokal nacheinander auf verschiedenen Tonhöhen hervorbrachte; war die Nachbildung des Klanges an dem Sprechapparat für eine bestimmte Tonhöhe richtig eingestellt, so brauchte man bei einer veränderten Tonhöhe nur die Frequenz des Stossoszillators entsprechend zu ändern, um wieder eine vollkommene Nachbildung des natürlichen Klanges zu erhalten. Es ist dabei aber nötig, die Mundstellung bei allen Tonhöhen unverändert zu lassen, was eine gewisse Übung erfordert.

Abb. 9 zeigt Schwingungsform und Spektrum des Vokals *i* beim Sprecher *V* auf der Tonhöhe von 275 Hertz für den Originalklang und die Nachbildung. Der Sprecher hat eine sehr charakteristisch helle Stimme: auch seine Vokale konnten naturgetreu nachgebildet werden. Die Formantwerte waren im Fall des Klanges nach Abb. 9

$$f_1 = 304 \text{ Hz}$$

$$d_1 = 0,18$$

$$f_2 = 2160 \text{ Hz}$$

$$d_2 = 0,07$$

$$f_3 = 3320 \text{ Hz}$$

$$d_3 = 0,043$$

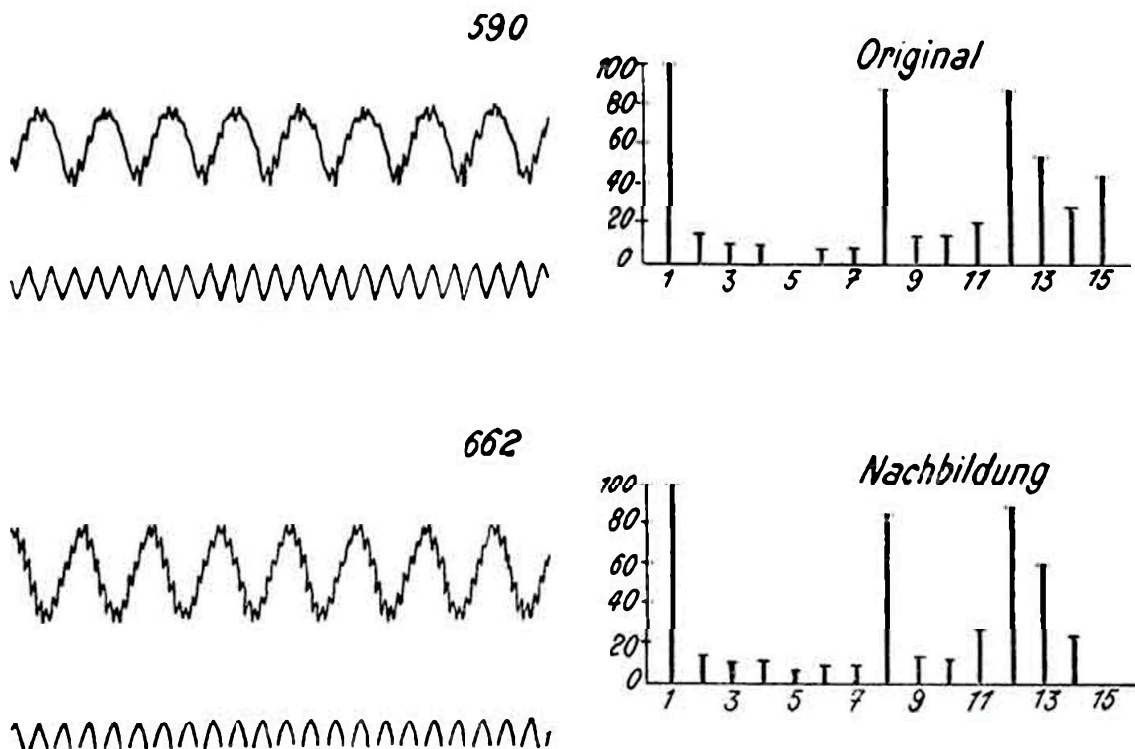


Abb. 9. Original und Nachbildung des Vokals *i* des Sprechers *V*,
Tonhöhe 275 Hz.

Abb. 10 veranschaulicht die Nachbildung des Vokals o desselben Sprechers bei der Tonhöhe 167 Hertz. Der Vokal hat den Hauptformant 350 Hz ($d = 1,0$) und einen kräftigen Nebenformant bei 690 Hz ($d = 0,23$).

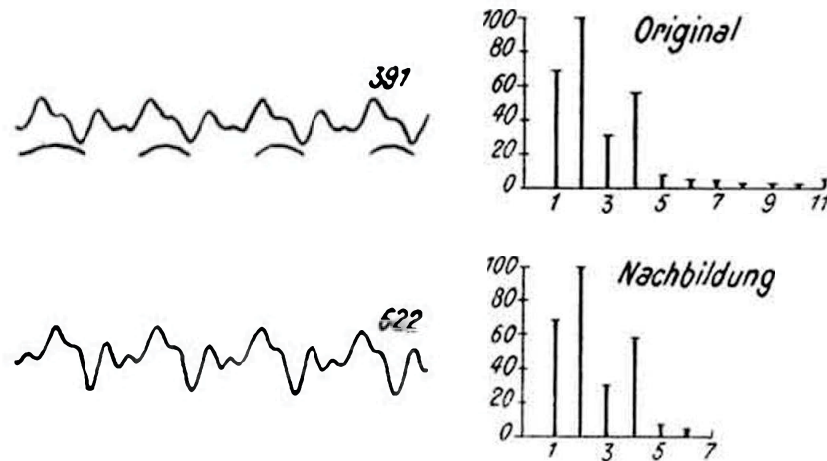


Abb. 10. Original und Nachbildung des Vokals o des Sprechers V, Tonhöhe 167 Hz.

Bei diesem Sprecher ändert sich die Klangfarbe mit der Tonhöhe sehr deutlich. Wir haben daher diese Erscheinung etwas eingehender untersucht und zeigen die Ergebnisse in Abb. 11. Bei tiefer Tonlage hat der Vokal einen Hauptformant bei etwa 400 Hz und ausserdem einen Nebenformant bei 600 Hz (Formantkurve 1). Mit steigender Höhe des Grundtones wird der Nebenformant immer schwächer (Formantkurve 2) und verschwindet beim Grundton 200 Hz vollständig (Formantkurve 3).

Die im Schrifttum mehrfach erörterte Frage, ob die Vokalformanten wirklich festliegen oder sich mit der Tonhöhe verschieben, ist wohl dahin zu beantworten, dass die Formanten sich mit der Tonhöhe nicht ändern, wenn die Klangfarbe des Vokals dieselbe bleibt. Das ist immer dann der Fall, wenn die Stellung des Sprechorgans bei einer Veränderung der Tonhöhe nicht mitgeändert wird. Bei vielen Personen ändert sich jedoch die Mundstellung, wenn ein bestimmter Vokal auf verschiedenen Tonhöhen in zwangloser Weise hervorgebracht wird; in diesem Falle ändert sich die Klangfarbe und verschieben sich die Formanten.

Mit dem hier beschriebenen Gerät können zahlreiche weitere Fragen über die Beschaffenheit von Vokalklängen aller Art untersucht werden, wie sie z. B. bei der Erforschung der verschiedenen

Dialekte und ihrer Bedingtheit durch körperliche Merkmale, bei pathologischen Veränderungen der Sprechorgane und in anderen Zusammenhängen auftreten.

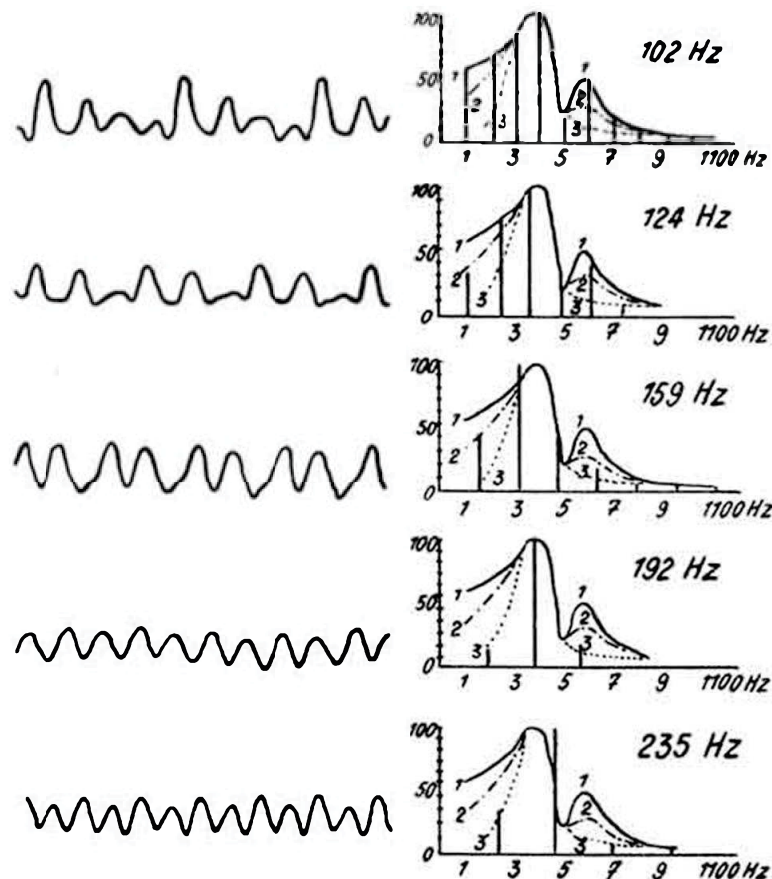


Abb. 11. Klangänderung mit der Tonhöhe durch Wegfall des Nebenformanten (Vokal o des Sprechers V)

Zusammenfassung

Die in dieser Arbeit beschriebene Einrichtung stellt eine elektrische Nachbildung des natürlichen Sprechorgans dar. Dem Kehlkopf, der eine regelmässige Folge von kurzen Luftstössen erzeugt, entspricht in der Nachbildung ein elektrischer Kreis, in dem entsprechende elektrische Stromstösse entstehen. Diese erregen eine Reihe von Schwingungskreisen, die sogenannten Formantkreise, deren Eigenfrequenz und Dämpfung mit den entsprechenden Eigenschaften der verschiedenen Hohlräume des Sprechorgans in Rachen, Mund und Nase übereinstimmen. Diese bestimmen ihrerseits den individuellen Vokalklang. Durch Veränderung der elektrischen Eigenschaften der Formantkreise kann man jeden Vokalklang einer

gegebenen Versuchsperson so genau nachbilden, dass der natürliche und der künstlich erzeugte Klang nicht unterscheidbar ist. Mit dieser Einrichtung lassen sich phonetische, physiologische und medizinische Untersuchungen über Sprachklänge ausführen.

Résumé

La méthode décrit l'imitation électro-acoustique de la production phonique. Grace à l'appareil se produisent des secousses régulières d'air qui provoquent une suite de vibrations dont la fréquence et l'amortissement correspondent aux qualités équivalentes des diverses cavités de résonance.

On peut déterminer le son individuel de la voyelle. En modifiant les capacités électriques on arrive à imiter le son de la voyelle particulier à chaque individu, au point qu'il est impossible de différencier le son naturel du son artificiel.

Cette technique est appelée à rendre de grands services en phonétique, en physiologie et en médecine.

Summary

The adjustment described herein displays an electrical imitation of the natural speech-organ. The larynx, producing a regular succession of short gusts of air is responded by an electrical circle in which corresponding impulses of current are formed. These cause a row of oscillation-circuits, the so-called formation circles, whose selffrequency and damping-circuit correspond with the responding character of the various hollow spaces of the speech-organs, in throat mouth and nose. These again determine the individual vowel-sound. By changing the electrical characteristic of the formation circuits every vowel-sound of the person to be tested can be imitated so accurately that no difference between the artificial and the natural sound can be noticed.

With this adjustment phonetical, physiological and medical investigations concerning speech-sounds, can be executed.